Unconstrained steepest descent method for multicriteria optimization on Riemmanian manifolds

G. C. Bento*   O. P. Ferreira †   P. R. Oliveira ‡

October 29, 2010

Abstract

In this paper we present a steepest descent method with Armijo's rule for multicriteria optimization in the Riemannian context. The well definedness of the sequence generated by the method is guaranteed. Under mild assumptions on the multicriteria function, we prove that each accumulation point (if they exist) satisfies first-order necessary conditions for Pareto optimality. Moreover, assuming quasi-convexity of the multicriteria function and non-negative curvature of the Riemannian manifold, we prove full convergence of the sequence to a Pareto critical.

Key words: Steepest descent, Pareto optimality, Vector optimization, Quasi-Fejér convergence, Quasi-convexity, Riemannian manifolds.

1 Introduction

Consider o following minimization problem

$$\begin{array}{l}\min F(p)\\ \text{s.t. } \; p \in X.\end{array} \tag{1}$$

In case that $F:\mathbb{R}^n \to \mathbb{R}$ and $X=\mathbb{R}^n$, the steepest descent method with Armijo's rule generates a sequence $\{p^k\}$ as follows

$$p^{k+1} = p^k + t_k v^k, \qquad v^k = -F'(p^k), \qquad k = 0, 1, \ldots,$$

where

$$t_k = \max\{2^{-j} : F(p^k - tv^k) \le F(p^k) + \beta v^k, j = 0, 1 \ldots\},$$

*IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, BR (Email: `glaydston@mat.ufg.br`)

†IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, BR (Email: `orizon@mat.ufg.br`). The author was supported in part by CNPq Grant 302618/2005-8, PRONEX–Optimization(FAPERJ/CNPq) and FUNAPE/UFG.

‡COPPE-Sistemas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21945-970, BR (Email: `poliveir@cos.ufrj.br`). This author was supported in part by CNPq.

$\beta \in (0,1)$. If F is continuously differentiable, classic results assure only that any accumulation point of $\{p^k\}$, case there exist, are critical of F. This fact was generalized for multicriteria optimization by Fliege and Svaiter [14], namely, whenever the objective function is a vectorial function $F : \mathbb{R}^n \to \mathbb{R}^m$ and the partial order in $\mathbb{R}^m$ is the usual, i.e., the componet-wise order. Full convergence is assured under the assumption that the solution set of the problem (1) is not-empty and $F : \mathbb{R}^n \to \mathbb{R}$ is a convex function, see Burachik et al. [6] (or, more generally, a quasi-convex function, see Kiwiel and Murty [19]), which has been generalized for vector optimization by Graña Drummond and Svaiter [17] (see also, Graña Drummond and Iusem [16]).

Extension of concepts, techniques as well as methods from Euclidean spaces to Riemannian manifolds is natural and, in general, nontrivial. In the last few years, such extension setting with purpose practical and theoretical has been the subject of many new research. Recent works dealing with this issue include [1, 2, 3, 4, 11, 8, 12, 21, 22, 23, 28, 30, 36, 37, 38]. The generalization of optimization methods from Euclidean space to Riemannian manifold have some important advantages. For example, constrained optimization problems can be seen as unconstrained one from the Riemannian geometry viewpoint (the set constrained is a manifold) and, in this case, we have an alternative possibility besides the projection idea for solving the problem. Moreover, nonconvex problems in the classical context may become convex through the introduction of an appropriate Riemannian metric (see, for example [8]).

The steepest descent method for the problem (1), in the particular case that $X = M$ (M a Riemannian manifold) and $F : M \to \mathbb{R}$ is continuously differentiable has been studied by Udriste [35], Smith [34] and Rapcsák [31] and partial convergence results were obtained. For the convex case the full convergence, using Armijo's rule, has been generalized by da Cruz Neto et al. [7], in the particular case that M has non-negative curvature. Regarding to the same restrictive assumption on the manifold M, Papa Quiroz et al. [28] generalized the full convergence result using generalized Armijo's rule for the quasiconvex case.

In this paper, following the ideas of Fliege and Svaiter [14], we generalize its converge results for multicriteria optimization to the Riemannian context. Besides, following the ideas of Graña Drummond and Svaiter [17], we generalize the full convergence result for multicriteria optimizationin in the case that the multicriteria function is quasi-convex and the Riemannian manifold has non-negative curvature.

The organization of our paper is as follows. In Section 2 we define the notations and list some results of Riemannian geometry to be used throughout this paper. In Section 3 we present the multicriteria problem, the first order optimality condition for it and some basic definitions related. In the Section 4 we state the Riemannian steepest descent methods for solving multicriteria problems and establish the well definition of the sequence generated for it. In Section 5 we prove a partial convergence result without any additional assumption on F besides the continuity differentiable and, assuming quasi-convexity of F and not-negative curvature for M, a full convergence result is presented. Finally, in Section 6 we present some examples of complete Riemannian manifolds with explicit geodesic curves and the steepest descent iteration of the sequence generated by the proposed method.

2 Preliminaries on Riemannian geometry

In this section, we introduce some fundamental properties and notations of Riemannian manifold. These basics facts can be found in any introductory book of Riemannian geometry, for example in [9] and [33].

Let M be a n-dimentional connected manifold. We denote by T_pM the n-dimensional *tangent space* of M at p, by $TM = \cup_{p\in M} T_pM$ *tangent bundle* of M and by $\mathcal{X}(M)$ the space of smooth vector fields over M. When M is endowed with a Riemannian metric $\langle\, ,\, \rangle$, with corresponding norm denoted by $\|\ \|$, then M is now a Riemannian manifold. Recall that the metric can be used to define the length of piecewise smooth curves $\gamma : [a, b] \to M$ joining p to q, i.e., such that $\gamma(a) = p$ and $\gamma(b) = q$, by

$$l(\gamma) = \int_a^b \|\gamma'(t)\| dt,$$

and, moreover, by minimizing this length functional over the set of all such curves, we obtain a Riemannian distance $d(p, q)$ which induces the original topology on M. The metric induces a map $f \mapsto \operatorname{grad} f \in \mathcal{X}(M)$ which associates to each scalar function smooth over M its gradient via the rule $\langle \operatorname{grad} f, X\rangle = df(X),\ X \in \mathcal{X}(M)$. Let ∇ be the Levi-Civita connection associated to $(M, \langle\, ,\, \rangle)$. A vector field V along γ is said to be *parallel* if $\nabla_{\gamma'} V = 0$. If γ' itself is parallel we say that γ is a *geodesic*. Because the geodesic equation $\nabla_{\ \gamma'}\gamma' = 0$ is a second order nonlinear ordinary differential equation, then the geodesic $\gamma = \gamma_v(., p)$ is determined by its position p and velocity v at p. It is easy to check that $\|\gamma'\|$ is constant. We say that γ is *normalized* if $\|\gamma'\| = 1$. The restriction of a geodesic to a closed bounded interval is called a *geodesic segment*. A geodesic segment joining p to q in M is said to be *minimal* if its length equals $d(p, q)$ and this geodesic is called a *minimizing geodesic*. If γ is a curve joining points p and q in M then, for each $t \in [a, b]$, ∇ induces an linear isometry, relative to $\langle\, ,\, \rangle$, $P_{\gamma(a)\gamma(t)} : T_{\gamma(a)}M \to T_{\gamma(t)}M$, the so-called *parallel transport* along γ from $\gamma(a)$ to $\gamma(t)$. The inverse map of $P_{\gamma(a)\gamma(t)}$ is denoted by $P^{-1}_{\gamma(a)\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(a)}M$. In the particular case of γ is the unique curve joining points p and q in M then parallel transport along γ from p to q is denoted by $P_{pq} : T_pM \to T_qM$.

A Riemannian manifold is *complete* if geodesics are defined for any values of t. Hopf-Rinow's theorem asserts that if this is the case then any pair of points, say p and q, in M can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, (M, d) is a complete metric space and bounded and closed subsets are compact. Take $p \in M$, the *exponential map* $exp_p : T_pM \to M$ is defined by $exp_p v = \gamma_v(1, p)$.

We denote by R *the curvature tensor* defined by $R(X, Y) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$, with $X, Y, Z \in \mathcal{X}(M)$, where $[X, Y] = YX - XY$. Then the *sectional curvature* with respect to X and Y is given by $K(X, Y) = \langle R(X, Y)Y, X\rangle/(||X||^2||Y||^2 - \langle X\ , Y\rangle^2)$, where $||X|| = \langle X, X\rangle^2$.

In the subsection 5.2 of this paper, we will be mainly interested in Riemannian manifolds for which $K(X, Y) \geq 0$ for any $X, Y \in \mathcal{X}(M)$. Such manifolds are referred to as *manifolds with nonnegative curvature*. A fundamental geometric property of this class of manifolds is that the

distance between geodesics issuing from one point is, at least locally, bounded from above by the distance between the corresponding rays in the tangent space. A global formulation of this general principle is the *law of cosines* that we now pass to describe. A *geodesic hinge* in M is a pair of normalized geodesics segment γ_1 and γ_2 such that $\gamma_1(0) = \gamma_2(0)$ and at least one of them, say γ_1, is minimal. From now on $l_1 = l(\gamma_1)$, $l_2 = l(\gamma_2)$, $l_3 = d(\gamma_1(l_1), \gamma_2(l_2))$ and $\alpha = \sphericalangle(\gamma_1'(0), \gamma_2'(0))$.

Theorem 1. *(Law of cosines) In a complete Riemannian manifold with nonnegative curvature, with the notation introduced above, we have*

$$l_3^2 \leq l_1^2 + l_2^2 - 2l_1 l_2 \cos \alpha. \tag{2}$$

Proof. See [9] and [33]. □

In this paper M will denote a complete n-dimensional Riemannian manifold.

3 The multicriteria problem

In this section we present the multicriteria problem, the first order optimality condition for it and some basic definitions related.

Let $I := \{1, \ldots, m\}$, $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x_i \geq 0, j \in I\}$ and $\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x_j > 0, j \in I\}$. For x, $y \in \mathbb{R}^m_+$, $y \succeq x$ (or $x \preceq y$) means that $y - x \in \mathbb{R}^m_+$ and $y \succ x$ (or $x \prec y$) means that $y - x \in \mathbb{R}^m_{++}$.

Given a vector continuously differentiable function $F : M \to \mathbb{R}^m$, we consider the problem of finding a *Pareto optimum point* of F, i.e., a point $p^* \in M$ such that there exists no other $p \in M$ with $F(p) \preceq F(p^*)$ and $F(p) \neq F(p^*)$. We denote this unconstrained problem in the Riemannian context as

$$\min_{p \in M} F(p). \tag{3}$$

Let F be given by $F(p) := (f_1(p), \ldots, f_m(p))$. We denote the Riemannian jacobian of F by

$$\operatorname{grad} F(p) := (\operatorname{grad} f_1(p), \ldots, \operatorname{grad} f_m(p)), \qquad p \in M,$$

and the image of the Riemannian jacobian of F at a point $p \in M$ by

$$\operatorname{Im}(\operatorname{grad} F(p)) := \{\operatorname{grad} F(p)v = (\langle \operatorname{grad} f_1(p), v\rangle, \ldots, \langle \operatorname{grad} f_m(p), v\rangle) : v \in T_pM\}, \qquad p \in M.$$

Using above equality the first order optimality condition for the problem (3) is stated as

$$p \in M, \qquad \operatorname{Im}(\operatorname{grad} F(p)) \cap (-\mathbb{R}^m_{++}) = \emptyset. \tag{4}$$

Remark 2. *Note that the condition in* (4) *generalizes to vector optimization the classical condition* $\operatorname{grad} F(p) = 0$ *for the scalar case, i.e.,* $m = 1$.

In general, (4) is necessary, but no sufficient for optimality. So, a point $p \in M$ satisfying (4) is called *Pareto critical.*

4 Steepest descent methods for multicriteria problems

In this section we state the Riemannian steepest descent methods for solving multicriteria problems and establish the well definition of the sequence generated for it.

Let $p \in M$ be a point which is not Pareto critical. Then there exists a direction $v \in T_pM$ satisfying

$$\operatorname{grad} F(p)v \in -\mathbb{R}^m_{++},$$

that is, $\operatorname{grad} F(p)v \prec 0$. In this case, v is called a *descent direction* for F at p.

For each $p \in M$, we consider the following unconstrained optimization problem in the tangent plane T_pM

$$\min_{v \in T_pM} \left\{ \max_{i\in I} \langle \operatorname{grad} f_i(p), v\rangle + (1/2)\|v\|^2 \right\}, \qquad I = \{1, \ldots, m\}. \tag{5}$$

Lemma 3. *The unconstrained optimization problem in* (5) *has only one solution. Moreover, the vector v is the solution of the problem in* (5) *if only if there exist $\alpha_i \geq 0$, $i \in I(p,v)$, such that*

$$v = -\sum_{i\in I(p,v)} \alpha_i \operatorname{grad} f_i(p), \qquad \sum_{i\in I(p,v)} \alpha_i = 1,$$

where $I(p,v) := \{i \in I : \langle \operatorname{grad} f_i(p), v\rangle = \max_{i\in I}\langle \operatorname{grad} f_i(p), v\rangle\}$.

Proof. Since the function

$$T_pM \ni v \mapsto \max_{i\in I}\langle \operatorname{grad} f_i(p), v\rangle,$$

is the maximum of linear functions in the linear space T_pM, it is convex. So, it is easy to see that the function

$$T_pM \ni v \longmapsto \max_{i\in I}\langle \operatorname{grad} f_i(p), v\rangle + (1/2)\|v\|^2, \tag{6}$$

is strong convex, which implies that the problem in (5) has only one solution in T_pM and the first statement is proved.

From convexity of the function in (6), it is well know that v is solution of the problem in (5) if only if

$$0 \in \partial \left(\max_{i\in I}\langle \operatorname{grad} f_i(p), \,.\,\rangle + (1/2)\|\,.\,\|^2\right)(v),$$

or equivalently,

$$-v \in \partial \left(\max_{i\in I}\langle \operatorname{grad} f_i(p), \,.\,\rangle\right)(v).$$

Therefore, the second statement follows of the formula for the subdifferential of the maximum of convex functions (see [18], Volume I, Corollary VI.4.3.2). $\square$

Lemma 4. *If $p \in M$ is not Pareto critical of F and v is the solution of the problem in* (5)*, then*

$$\max_{i\in I}\langle \operatorname{grad} f_i(p), v\rangle + (1/2)\|v\|^2 < 0.$$

In particular, v is a descent direction.

Proof. Since p is not Pareto critical, there exists $0 \neq \hat{v} \in T_pM$ such that $\operatorname{grad} F(p)\hat{v} \prec 0$. In particular,

$$\beta = \max_{i\in I}\langle \operatorname{grad} f_i(p), \hat{v}\rangle < 0.$$

As $-\beta/\|\hat{v}\|^2 > 0$, letting $\bar{v} = (-\beta/\|\hat{v}\|^2)\hat{v}$ we obtain

$$\max_{i\in I}\langle \operatorname{grad} f_i(p), \bar{v}\rangle + (1/2)\|\bar{v}\|^2 = -\frac{\beta^2}{2\|\hat{v}\|^2} < 0,$$

Using that v is the solution of the problem in (5), the first part of the lemma follows from last inequality. The second part of the lemma is an immediate consequence of the first one. □

In view of the two previous lemmas and (5) we define the steepest descent direction function for F as follows.

Definition 5. *The steepest descent direction function for F is defined as*

$$M \ni p \longmapsto v(p) := \operatorname{argmin}_{v\in T_pM} \left\{\max_{i\in I} \langle \operatorname{grad} f_i(p), v\rangle + (1/2)\|v\|^2\right\} \in T_pM.$$

Remark 6. *As an immediate consequence of Lemma 3 it follows that the steepest descent direction for vector functions becomes the steepest descent direction when $m = 1$. See, for example, [7], [26], [31], [34] and [35]. In the case $M = \mathbb{R}^n$ we retrieve the steepest descent direction proposed in [14].*

The *steepest descent method with Armijo rule* for solving the unconstrained optimization problem (3) is as follows:

Method 1 (Steepest descent method with Armijo rule)**.**

Initialization. *Take $\beta \in (0, 1)$ and $p_0 \in M$. Set $k = 0$.*
Stop criterion. *If p^k is Pareto critical STOP. Otherwise.*
Iterative Step. *Compute the steepest descent direction v^k for F at p^k, i.e.,*

$$v^k := v(p^k), \tag{7}$$

and the steplength $t_k \in (0, 1]$ is of the following way:

$$t_k := \max\left\{2^{-j} : j \in \mathbb{N},\ F\left(exp_{p^k}(2^{-j}v^k)\right) \preceq F(p^k) + \beta 2^{-j} \operatorname{grad} F(p^k)v^k\right\}, \tag{8}$$

and set

$$p^{k+1} := exp_{p^k}(t_k v^k), \tag{9}$$

and GOTO Stop criterion.

Remark 7. *The steepest descent method for vector optimization in Riemannian manifolds becomes the classical steepest descent method when $m = 1$, which has appeared, for example, in [7], [31], [34] and [35].*

Proposition 8. *The sequence $\{p^k\}$ generated by steepest descent method with Armijo rule is well defined.*

Proof. Assume that p^k is not Pareto critical. From Definition 5 and Lemma 3, $v^k = v(p^k)$ is well defined. Thus, for proving the well definition of the method proposed it is enough proving well definition of the steplength. For this, first note that from Definition 5 and Lemma 4

$$\operatorname{grad} F(p^k)v^k \prec 0.$$

Since $F : M \to \mathbb{R}^m$ is a continuously differentiable vector function, $\operatorname{grad} F(p^k)v^k \prec 0$ and $\beta \in (0,1)$ we have

$$\lim_{t\to 0^+} \frac{F\left(exp_{p^k}(tv^k)\right) - F(p^k)}{t} = \operatorname{grad} F(p^k)v^k \prec \beta \operatorname{grad} F(p^k)v^k \prec 0.$$

Therefore, it is straightforward to show that there exists $\delta \in (0,1]$ such that

$$F\left(exp_{p^k}(tv^k)\right) \prec F(p^k) + \beta t \operatorname{grad} F(p^k)v^k, \qquad t \in (0,\, \delta).$$

As $\lim_{j\to\infty} 2^{-j} = 0$, last vector inequality implies that the steplength (8) is well defined. Hence p^{k+1} is also well defined and the proposition is concluded. □

5 Convergence analysis

In this section, following the ideas of [14] we prove a partial convergence result without any additional assumption on F besides the continuity differentiable. In the sequel, following [17], assuming quasi-convexity of F and not-negative curvature for M, we extend to optimization of vector functions the full convergence result presented in [7] and [28]. It is immediate to see that, if the Method 1 terminates after a finite number of iterations, it terminates at a Pareto critical point. From now on, we will assume that $\{p^k\}$, $\{v^k\}$ and $\{t_k\}$ are infinite sequences generated by Method 1.

5.1 Partial convergence result

In this subsection we prove that every accumulation point of $\{p^k\}$ is a Pareto critical point. Before this, we prove the following preliminary fact that will be useful.

Lemma 9. *The steepest descent direction function for F, $M \ni p \mapsto v(p) \in T_pM$, is continuous.*

Proof. Let $\{q^k\} \subset M$ be a sequence which converges to $\bar{q}$ as k goes to $+\infty$, and $U_{\bar{q}} \subset M$ a neighborhood of $\bar{q}$ such that $TU_{\bar{q}} \approx U_{\bar{q}} \times \mathbb{R}^n$. Since $\{q^k\}$ converges to $\bar{q}$ and $TU_{\bar{q}} \subset TM$ is an open set, we assume that the whole sequence $\{(q^k, v(q^k))\}$ is in $TU_{\bar{q}}$. Define $v^k := v(q^k)$. Combining Definition 5 with Lemma 4 it is easy to see that

$$\|v^k\| \leq 2\max_{i\in I} \|\operatorname{grad} f_j(q^k)\|.$$

As F is continuously differentiable and $\{q^k\}$ is convergent, above inequality implies that the sequence $\{v^k\}$ is bounded. Let $\bar{v}$ be an accumulation point of the sequence $\{v^k\}$. From Definition 5 and Lemma 3 we conclude that there exist $\alpha_i^k \geq 0$, $i \in I(q^k, v^k)$, such that

$$v^k = -\sum_{i \in I(q^k, v^k)} \alpha_i^k \operatorname{grad} f_i(q^k), \qquad \sum_{i \in I(q^k, v^k)} \alpha_i^k = 1, \qquad k = 0, 1, \ldots. \tag{10}$$

where $I(q^k, v^k) := \{i \in I : \langle \operatorname{grad} f_i(q^k), v^k \rangle = \max_{i \in I} \langle \operatorname{grad} f_i(q^k), v^k \rangle\}$. Using above constants and the associated indexes, define the sequence $\{\alpha^k\}$ as

$$\alpha^k := (\alpha_1^k, \ldots, \alpha_m^k), \qquad \alpha_i^k = 0, \quad i \in I \setminus I(q^k, v^k), \qquad k = 0, 1, \ldots.$$

Let $\|\,.\,\|_1$ be the sum norm in $\mathbb{R}^m$. Since $\sum_{i \in I(q^k, v^k)} \alpha_i^k = 1$, we have $\|\alpha^k\|_1 = 1$ for all k, which implies that the sequence $\{\alpha^k\}$ is bounded. Let $\bar{\alpha}$ be an accumulation point of the sequence $\{\alpha^k\}$. Let $\{v^{k_s}\}$ and $\{\alpha^{k_s}\}$ be subsequences of $\{v^k\}$ and $\{\alpha^k\}$ respectively, such that

$$\lim_{s \to +\infty} v^{k_s} = \bar{v}, \qquad \lim_{s \to +\infty} \alpha^{k_s} = \bar{\alpha}.$$

As the index set I is finite and $I(q^{k_s}, v^{k_s}) \subset I$ for all s, we assume without loss of generality that

$$I(q^{k_1}, v^{k_1}) = I(q^{k_2}, v^{k_2}) = ... = \bar{I}. \tag{11}$$

Hence, we conclude from (10) and last equalities that

$$v^{k_s} = -\sum_{i \in \bar{I}} \alpha_i^{k_s} \operatorname{grad} f_i(q^{k_s}), \qquad \sum_{i \in \bar{I}} \alpha_i^{k_s} = 1, \qquad s = 0, 1, \ldots.$$

Letting s goes to $+\infty$ in the above equalities, we obtain

$$\bar{v} = \sum_{i \in \bar{I}} \bar{\alpha}_i \operatorname{grad} f_i(\bar{q}), \qquad \sum_{i \in \bar{I}} \bar{\alpha}_i = 1. \tag{12}$$

On the other hand, $I(q^{k_s}, v^{k_s}) = \{i \in I : \langle \operatorname{grad} f_i(q^{k_s}), v^{k_s} \rangle = \max_{i \in I} \langle \operatorname{grad} f_i(q^{k_s}), v^{k_s} \rangle\}$. So, equation (11) implies that

$$\langle \operatorname{grad} f_i(q^{k_s}), v^{k_s} \rangle = \max_{i \in I} \langle \operatorname{grad} f_i(q^{k_s}), v^{k_s} \rangle, \qquad i \in \bar{I}, \qquad s = 0, 1, \ldots.$$

Using continuity of $\operatorname{grad} F$ and last equality we have

$$\langle \operatorname{grad} f_i(\bar{q}), \bar{v} \rangle = \max_{i \in I} \langle \operatorname{grad} f_i(\bar{q}), \bar{v} \rangle, \qquad i \in \bar{I}.$$

From the definition of $I(\bar{q}, \bar{v})$ we obtain $\bar{I} \subset I(\bar{q}, \bar{v})$. Therefore, combining again Definition 5 with Lemma 3 and (12), we conclude that $\bar{v} = v(\bar{q})$ and the desired result is proved. $\square$

In the next result, we just use that F is continuously differentiable to assure that the sequence of the functional values of the sequence $\{p^k\}$, $\{F(p^k)\}$, it is monotonous decreasing and that their accumulation points are critical Pareto.

Theorem 10. *The following statements there hold:*

i) $\{F(p^k)\}$ *is decreasing;*

ii) Each accumulation point of the sequence $\{p^k\}$ *is a Pareto critical point.*

Proof. The iterative step in the Method 1 implies that

$$F(p^{k+1}) \preceq F(p^k) + \beta t_k \operatorname{grad} F(p^k) v^k, \qquad p^{k+1} = \exp_{p^k} t_k v^k, \qquad k = 0, 1, \ldots. \tag{13}$$

Since $\{p^k\}$ is a infinite sequence, for all k, p^k is not Pareto critical of F. Thus, item i follows from definition of v^k together with Definition 5, Lemma 4 and last vector inequality.

Let $\bar{p} \in M$ be an accumulation point of the sequence $\{p^k\}$ and $\{p^{k_s}\}$ a subsequence of $\{p^k\}$ such that $\lim_{s\to+\infty} p^{k_s} = \bar{p}$. Since F is continuous and $\lim_{s\to+\infty} p^{k_s} = \bar{p}$ we have $\lim_{s\to+\infty} F(p^{k_s}) = F(\bar{p})$. So, taking into account that $\{F(p^k)\}$ is a decreasing sequence and has $F(\bar{p})$ as an accumulation point, it is easy to conclude that the whole sequence $\{F(p^k)\}$ converges to $F(\bar{p})$. Using the equation (13), Definition 5 and Lemma 4, we conclude that

$$F(p^{k+1}) - F(p^k) \preceq \beta t_k \operatorname{grad} F(p^k) v^k \preceq 0, \qquad k = 0, 1, \ldots.$$

Since $\lim_{s\to+\infty} F(p^k) = F(\bar{p})$, last inequality implies that

$$\lim_{k\to+\infty} \beta t_k \operatorname{grad} F(p^k) v^k = 0. \tag{14}$$

As $\{p^{k_s}\}$ converges to $\bar{p}$, we assume that $\{(p^{k_s}, v^{k_s})\} \subset TU_{\bar{p}}$, where $U_{\bar{p}}$ is a neighborhood of $\bar{p}$ such that $TU_{\bar{p}} \approx U_{\bar{p}} \times \mathbb{R}^n$. Moreover, as the sequence $\{t_k\} \subset (0, 1]$ has an accumulation point $\bar{t} \in [0, 1]$, we assume without loss of generality that $\{t_{k_s}\}$ converges to $\bar{t}$. We have two possibilities to consider:

a) $\bar{t} > 0$;

b) $\bar{t} = 0$.

Assume that the item **a** holds. In this case, from (14), continuity of $\operatorname{grad} F$, (7) and Lemma 9, we obtain

$$\operatorname{grad} F(\bar{p}) v(\bar{p}) = 0,$$

which implies that

$$\max_{i\in I} \langle \operatorname{grad} f_i(\bar{p}), v(\bar{p}) \rangle = 0. \tag{15}$$

On the other hand, from Definition 5 together with Lemma 4,

$$\max_{i\in I} \langle \operatorname{grad} f_i(p^{k_s}), v^{k_s} \rangle + (1/2)\|v^{k_s}\|^2 < 0.$$

Letting s goes to $+\infty$ in the above inequalities and using Lemma 9 combined with the continuity of $\operatorname{grad} F$ and equality (15), we conclude that

$$\max_{i\in I}\langle \operatorname{grad} f_i(\bar p), v(\bar p)\rangle + (1/2)\|v(\bar p)\|^2 = 0.$$

Hence, it follows from last equality, Definition 5 and Lemma 4 that $\bar p$ is a Pareto critical.

Now, assume that the item **b** holds. Since for all s p^{k_s} is not a Pareto critical, we have

$$\max_{i\in I}\langle \operatorname{grad} f_i(p^{k_s}), v^{k_s}\rangle \le \max_{i\in I}\langle \operatorname{grad} f_i(p^{k_s}), v^{k_s}\rangle + (1/2)\|v^{k_s}\|^2 < 0,$$

where the last inequality is consequence from Definition 5 together with Lemma 4. Hence, letting s goes to $+\infty$ in the last inequalities, using (7) and Lemma 9 we obtain

$$\max_{i\in I}\langle \operatorname{grad} f_i(\bar p), v(\bar p)\rangle \le \max_{i\in I}\langle \operatorname{grad} f_i(\bar p), v(\bar p)\rangle + (1/2)\|v(\bar p)\|^2 \le 0. \tag{16}$$

Take $r \in \mathbb{N}$. Since $\{t_{k_s}\}$ converges to $\bar t = 0$, we conclude that for s large enough,

$$t_{k_s} < 2^{-r}.$$

From (8) this means that the Armijo condition (13) is not satisfied for $t = 2^{-r}$, i.e.,

$$F(\exp_{p^k}(2^{-j}v^{k_s})) \npreceq F(p^{k_s}) + \beta 2^{-r} \operatorname{grad} F(p^{k_s})v^{k_s},$$

which means that there exists at least one $i_0 \in I$ such that

$$f_{i_0}(\exp_{p^{k_s}}(2^{-r}v^{k_s})) > f_{i_0}(p^{k_s}) + \beta 2^{-r}\langle \operatorname{grad} f_{i_0}(p^{k_s}), v^{k_s}\rangle.$$

Letting s goes to $+\infty$ in the above inequality, taking into account that $\operatorname{grad} F$ and exp are continuous and using Lemma 9, we obtain

$$f_{i_0}(\exp_{\bar p}(2^{-r}v(\bar p))) \ge f_{i_0}(\bar p) + \beta 2^{-r}\langle \operatorname{grad} f_{i_0}(\bar p), v(\bar p)\rangle.$$

Last inequality is equivalent to

$$\frac{f_{i_0}(\exp_{\bar p}(2^{-r}v(\bar p))) - f_{i_0}(\bar p)}{2^{-r}} \ge \beta\langle \operatorname{grad} f_{i_0}(\bar p), v(\bar p)\rangle,$$

which, letting r goes to $+\infty$ and using that $0 < \beta < 1$ yields $\langle \operatorname{grad} f_{i_0}(\bar p), v(\bar p)\rangle \ge 0$. Hence,

$$\max_{i\in I}\langle \operatorname{grad} f_i(\bar p), v(\bar p)\rangle \ge 0.$$

Combining last inequality with (16), we have

$$\max_{i\in I}\langle \operatorname{grad} f_i(\bar p), v(\bar p)\rangle + (1/2)\|v(\bar p)\|^2 = 0.$$

Therefore, again from Definition 5 and Lemma 4 it follows that $\bar p$ is a Pareto critical and the proof is concluded. $\square$

Remark 11. *If the sequence $\{p^k\}$ begins in a bounded level set, for example, if*

$$L_F(F(p_0)) := \{p \in M : F(p) \preceq F(p_0)\},$$

is a bounded set, and being F a continuous function, Hopf-Rinow's theorem assures that $L_F(F(p_0))$ is a compact set. So, item i of Theorem 10 implies that $\{p^k\} \subset L_F(F(p_0))$ and consequently $\{p^k\}$ is bounded. In particular, $\{p^k\}$ has at least one accumulation point. Therefore, Theorem 10 extends for vector optimization the results of Theorem 5.1 *of [7]. See also Remark* 4.5 *of [15].*

5.2 Full convergence

In this section under the quasi-convexity assumption on F and not-negative curvature for M, full convergence of the steepest descent method is obtained.

Definition 12. *Let $H : M \to \mathbb{R}^m$ be a vectorial function.*

i) H is called convex on M if for every $p, q \in M$ and every geodesic segment $\gamma : [0,1] \to M$ joining p to q (i.e., $\gamma(0) = p$ and $\gamma(1) = q$), it holds

$$H(\gamma(t)) \preceq (1-t)H(p) + tH(q), \qquad t \in [0,1].$$

ii) H is called quasi-convex on M if for every $p, q \in M$ and every geodesic segment $\gamma : [0,1] \to M$ joining p to q, it holds

$$H(\gamma(t)) \preceq \max\{H(p), H(q)\}, \qquad t \in [0,1],$$

where the maximum is considered coordinate by coordinate.

Remark 13. *The first above definition is a natural extension of the definition of convexity while the second is an extension of a characterization of the definition of quasi-convexity, of the Euclidean space to the Riemannian context. See Definition* 6.2 *and Corollary* 6.6 *of [24], pages* 29 *and* 31 *respectively. Note that the above definitions are equivalent, respectively, H to be convex and quasi-convex along every geodesic segment. Thus, when $m = 1$ these definitions merge into the scalar convexity and quasi-convexity defined in [35], respectively. Moreover, it is immediate of the above definitions that if H is convex then it is quasi-convex. In the case that H is differentiable, convexity of H implies that for every $p, q \in M$ and every geodesic segment $\gamma : [0,1] \to M$ such that $\gamma(0) = p$ and $\gamma(1) = q$,*

$$\operatorname{grad} H(p)\gamma'(0) \preceq H(q) - H(p).$$

Proposition 14. *Let $H : M \to \mathbb{R}^m$ be a differentiable quasi-convex function. Then, for every $p, q \in M$ and every geodesic segment $\gamma : [0,1] \to M$ joining p to q, it holds*

$$H(q) \preceq H(p) \quad \Rightarrow \quad \operatorname{grad} H(p)\gamma'(0) \preceq 0.$$

Proof. Take $p, q \in M$ such that $H(q) \preceq H(p)$ and a geodesic segment $\gamma : [0,1] \to M$ such that $\gamma(0) = p$ and $\gamma(1) = q$. Since H is quasi-convex, we have

$$H(\gamma(t)) \preceq H(p), \qquad t \in [0,1].$$

Using last inequality the result is an immediate consequence from the differentiability of H. $\square$

We know that criticality is necessary condition but not sufficient for optimality. However, under convexity of the vectorial function F we proved that criticality is equivalent to the weak optimality.

Definition 15. *A point $p^* \in M$ is a weak Pareto optimal of F if there is no $p \in M$ with $F(p) \prec F(p^*)$.*

Proposition 16. *Let $H : M \to \mathbb{R}^m$ be a convex continuously differentiable function. Then, $p \in M$ is a Pareto critical of H, i.e.,*

$$Im(\operatorname{grad} H(p)) \cap (-\mathbb{R}^m_{++}) = \emptyset,$$

if and only if p is a weak Pareto optimal of H.

Proof. Let us suppose that p is Pareto critical of H. Assume by contradiction that p is not weak Pareto optimal of H. Since p is not weak Pareto optimal, there exists $\tilde{p} \in M$ such that

$$H(\tilde{p}) \prec H(p). \tag{17}$$

Let $\gamma : [0,1] \to M$ be a geodesic segment joining p to $\tilde{p}$ (i.e., $\gamma(0) = p$ and $\gamma(1) = \tilde{p}$). As H is differentiable and convex, the last part of Remark 13 and (17) imply that

$$\operatorname{grad} H(p)\gamma'(0) \preceq H(\tilde{p}) - H(p) \prec 0.$$

But this contradicts the fact of p to be Pareto critical of H, and the first part is concluded.

Now, let us suppose that p is weak Pareto optimal of H. Assume by contradiction that p is not Pareto critical of H. Since p is not Pareto critical, then $\text{Im}(\operatorname{grad} H(p)) \cap (-\mathbb{R}^m_{++}) \neq \emptyset$, that is, there exists $v \in T_pM$ a descent direction for F at p. Hence, from the differentiability of H, we have

$$\lim_{t \to 0^+} \frac{H\left(exp_p(tv)\right) - H(p)}{t} = \operatorname{grad} H(p)v \prec 0,$$

which implies that there exists $\delta > 0$ such that

$$H\left(exp_p(tv)\right) \prec H(p) + t \operatorname{grad} H(p)v, \qquad t \in (0,\, \delta).$$

Since v is a descent direction for F at p and $t \in (0,\, \delta)$ we have $t \operatorname{grad} H(p)v \prec 0$. So, the last vector inequality yields

$$H\left(exp_p(tv)\right) \prec H(p), \qquad t \in (0,\, \delta),$$

contradicting the fact of p to be weak Pareto optimal of H, which concludes the proof. $\square$

Definition 17. *A sequence $\{q^k\} \subset M$ is quasi-Fejér convergent to a nonempty set U if, for all $p \in U$, there exists a sequence $\{\epsilon_k\} \subset \mathbb{R}_+$ such that*

$$\sum_{k=0}^{+\infty} \epsilon_k < +\infty, \qquad d^2(q^{k+1}, q) \leq d^2(q^k, q) + \epsilon_k, \qquad k = 0, 1, \ldots.$$

In the next lemma we recall the called quasi-Fejér convergence theorem.

Lemma 18. *Let $U \subset M$ be a nonempty set and $\{q^k\} \subset M$ a sequence quasi-Fejér convergent. Then, $\{q^k\}$ is bounded. Moreover, if an accumulation point $\bar{q}$ of $\{q^k\}$ belongs to U, then the whole sequence $\{q^k\}$ converges to $\bar{q}$ as k goes to $+\infty$.*

Proof. Analogous to the proof of Theorem 1 in Burachik et al. [6], by replacing the Euclidean distance by the Riemannian distance d. □

Consider the following set

$$U := \{p \in M : F(p) \preceq F(p^k), \;\; k = 0, 1, \ldots\}. \tag{18}$$

In general, the above set may be an empty set. To guarantee that U is nonempty, an additional assumption on the sequence $\{p^k\}$ is needed. In the next remark we give a such condition.

Remark 19. *If the sequence $\{p^k\}$ has an accumulation point, then U is nonempty. Indeed, let $\bar{p}$ be an accumulation point of the sequence $\{p^k\}$. Then, there exists a subsequence $\{p^{k_j}\}$ of $\{p^k\}$ which converges to $\bar{p}$. Since F is continuous $\{F(p^k)\}$ has $F(\bar{p})$ as an accumulation point. Hence, using that $\{F(p^k)\}$ is a decreasing sequence (see item i of Theorem 10) usual arguments show easily that the whole sequence $\{F(p^k)\}$ converges to $F(\bar{p})$ and there holds*

$$F(\bar{p}) \preceq F(p^k), \qquad k = 0, 1, \ldots,$$

which implies that $\bar{p} \in U$, i.e., $U \neq \emptyset$.

In the next lemma we presented the main result of this section. It is fundamental in the proof of the global convergence result of the sequence $\{p^k\}$.

Lemma 20. *Suppose that F is quasi-convex, M has not-negative curvature and U, defined in* (18), *is nonempty. Then, for all $\tilde{p} \in U$, the inequality there holds:*

$$d^2(p^{k+1}, \tilde{p}) \leq d^2(p^k, \tilde{p}) + t_k^2 \|v^k\|^2.$$

Proof. Consider the geodesic hinge $(\gamma_1, \gamma_2, \alpha)$, where γ_1 is a normalized minimal geodesic segment joining p^k to $\tilde{p}$, γ_2 is the geodesic segment joining p^k to p^{k+1} such that $\gamma_2'(0) = t_k v^k$ and $\alpha = \angle(\gamma_1'(0), v^k)$. By the law of cosines (Theorem 1), we have

$$d^2(p^{k+1}, \tilde{p}) \leq d^2(p^k, \tilde{p}) + t_k^2 \|v^k\|^2 - 2d(p^k, p) t_k \|v^k\| \cos\alpha, \qquad k = 0, 1, \ldots.$$

Thus, taking into account that $\cos(\pi - \alpha) = -\cos\alpha$ and $\langle -v^k, \gamma_1'(0)\rangle = \|v^k\| \cos(\pi - \alpha)$, above vector inequality becomes

$$d^2(p^{k+1}, \tilde{p}) \leq d^2(p^k, \tilde{p}) + 2d(p^k, \tilde{p}) t_k \langle -v^k, \gamma_1'(0)\rangle, \qquad k = 0, 1, \ldots.$$

On the other hand, from (7), Definition 5 and Lemma 3, there exist $\alpha_i^k \geq 0$, with $i \in I_k := I(p^k, v^k)$, such that

$$v^k = -\sum_{i \in I_k} \alpha_i \operatorname{grad} f_i(p^k), \qquad \sum_{i \in I_k} \alpha_i^k = 1, \qquad k = 0, 1, \ldots.$$

Hence, last vector inequality yields

$$d^2(p^{k+1}, \tilde{p}) \leq d^2(p^k, \tilde{p}) + 2d(p^k, \tilde{p}) t_k \sum_{i \in I_k} \alpha_i^k \langle \operatorname{grad} f_i(p^k), \gamma_1'(0)\rangle, \quad k = 0, 1, \ldots. \tag{19}$$

Since F is quasi-convex and $\tilde{p} \in U$, from Proposition 14 with $H = F$, $p = p^k$, $q = \tilde{p}$ and $\gamma = \gamma_1$, we have

$$\operatorname{grad} F(p^k) \gamma_1'(0) \preceq 0, \qquad k = 0, 1, \ldots,$$

or equivalently,

$$\langle \operatorname{grad} f_i(p^k), \gamma_1'(0)\rangle \leq 0, \qquad i = 1, \ldots, m, \quad k = 0, 1, \ldots. \tag{20}$$

Therefore, for combining (19) with (20), the lemma follows. □

Proposition 21. *If F is quasi-convex, M has not-negative curvature and U, defined in* (18)*, is a nonempty set, then the sequence $\{p_k\}$ is Fejér convergent to U.*

Proof. For simplify the notation define the scalar function $\varphi : \mathbb{R}^m \to \mathbb{R}$ as follows

$$\varphi(y) = \max_{i \in I} \langle y, e_i\rangle, \qquad I = \{1, \ldots, m\}.$$

where $\{e_i\} \subset \mathbb{R}^m$ is the canonical base of the space $\mathbb{R}^m$. It easy to see that the following properties on the function φ hold:

$$\varphi(x + y) \leq \varphi(x) + \varphi(y), \qquad \varphi(tx) = t\varphi(x), \qquad x, y \in \mathbb{R}^m, \quad t \geq 0. \tag{21}$$

$$x \preceq y \quad \Rightarrow \quad \varphi(x) \leq \varphi(y), \qquad x, y \in \mathbb{R}^m. \tag{22}$$

From the definition of t_k in (8) and p^{k+1} in (9), we have

$$F(p^{k+1}) \preceq F(p^k) + \beta t_k \operatorname{grad} F(p^k) v^k, \qquad k = 0, 1 \ldots.$$

Hence, using (21), (22) and last inequality, we obtain

$$\varphi(F(p^{k+1})) \leq \varphi(F(p^k)) + \beta t_k \varphi(\operatorname{grad} F(p^k) v^k), \qquad k = 0, 1 \ldots. \tag{23}$$

On the other hand, combining definition of v^k in (7), Definition 5, Lemma 4 and definition of φ, we conclude that

$$\varphi(\operatorname{grad} F(p^k)v^k) + (1/2)\|v^k\|^2 < 0, \qquad k = 0, 1\ldots,$$

which together with (23) implies that

$$\varphi(F(p^{k+1})) < \varphi(F(p^k)) - (\beta t_k/2)\|v^k\|^2, \qquad k = 0, 1\ldots,$$

But this tells us that,

$$t_k\|v^k\|^2 < 2[\varphi(F(p^k)) - \varphi(F(p^{k+1}))]/\beta, \qquad k = 0, 1\ldots.$$

As $t_k \in (0, 1]$, follows that

$$t_k^2\|v^k\|^2 < 2[\varphi(F(p^k)) - \varphi(F(p^{k+1}))]/\beta, \qquad k = 0, 1\ldots.$$

Thus, the latter inequality implies easily that

$$\sum_{k=0}^{n} t_k^2\|v^k\|^2 < 2\left[\varphi(F(p^0) - \varphi(F(p^{n+1}))\right]/\beta, \quad n > 0.$$

Take $\bar{p} \in U$. Then, $F(\bar{p}) \preceq F(p^{n+1})$. So, from (22) $\varphi(F(\bar{p})) \leq \varphi(F(p^{n+1}))$ and last inequality yields

$$\sum_{k=0}^{n} t_k^2\|v^k\|^2 < 2\left(\varphi(F(p_0) - \varphi(F(\bar{p}))\right)/\beta.$$

which implies that $\{t_k^2\|v^k\|^2\}$ is a summable sequence. Therefore, the desired result follows from Lemma 20 combined with Definition 17. □

Theorem 22. *If F is quasi-convex, M has not-negative curvature and U, defined in* (18)*, is a nonempty set, then the sequence $\{p_k\}$ converges to a Pareto critical of F.*

Proof. From Proposition 21, $\{p^k\}$ is Fejér convergent to U. Thus Lemma 18 guarantees that $\{p^k\}$ is bounded and, from Hopf-Rinow' theorem, there exists $\{p^{k_s}\}$, subsequence of $\{p^k\}$, which converges to $\bar{p} \in M$ as s goes to $+\infty$. Since F is continuous and $\{F(p^k)\}$ is a decreasing sequence (see item i of Theorem 10), we conclude that $F(p^k)$ converges to $F(\bar{p})$ as k goes to $+\infty$, which implies that

$$F(\bar{p}) \preceq F(p^k), \qquad k = 0, 1, \ldots,$$

i.e., $\bar{p} \in U$. Hence, from Lemma 18 we conclude that the whole sequence $\{p^k\}$ converges to $\bar{p}$ as k goes to $+\infty$, and the conclusion of the proof it is consequence of the item ii of Theorem 10. □

Corollary 23. *If F is convex, M has not-negative curvature and U, defined in* (18)*, is a nonempty set, then the sequence $\{p_k\}$ converges to a weak Pareto optimal of F.*

Proof. Since F is convex, in particular, it is quasi-convex (see Remark 13). Thus the corollary is a consequence of the previous theorem and Proposition 16. □

6 Examples

In this section we present some examples of complete Riemannian manifolds with explicit geodesic curves and the steepest descent iteration of the sequence generated by the method 1. We recall that $F : M \to \mathbb{R}^m$, $F(p) := (f_1(p), \ldots, f_m(p))$, is a differentiable function. If (M, G) is a Riemannian manifold then the Riemannian gradient of f_i is given by $\operatorname{grad} f_i(p) = G(p)^{-1} f_i'(p)$, $i \in I := \{1, \ldots, n\}$. Hence, if $v(p)$ is the steepest descent direction for F at p (see Definition 5), from Lemma 3, there exist constants $\alpha_i \geq 0$, $i \in I(p, v)$, such that

$$v = -\sum_{i \in I(p,v)} \alpha_i G(p)^{-1} f_i'(p), \qquad \sum_{i \in I(p,v)} \alpha_i = 1, \tag{24}$$

where $I(p, v) := \{i \in I : \langle G(p)^{-1} f_i'(p), v\rangle = \max_{i \in I} \langle G(p)^{-1} f_i'(p), v\rangle\}$.

6.1 A steepest descent method for $\mathbb{R}^n_{++}$

Let M be the positive octant, $\mathbb{R}^n_{++}$, endowed with the Riemannian metric

$$M \ni v \mapsto G(p) = P^{-2} := \operatorname{diag}\left(p_1^{-2}, \ldots, p_n^{-2}\right),$$

(metric induzed by the Hessian of the logarithmic barrier). Since (M, G) is isometric to the Euclidean space endowed with the usual metric (see, Da Cruz Neto et al. [8]) it follows that M has constant curvature equal to zero. On the other hand, it is easy to see that the unique geodesic $p = p(t)$ such that $p(0) = p^0 = (p_1^0, \ldots, p_n^0)$ and $p'(0) = v^0 = (v_1^0, \ldots, v_n^0)$ is given by $p(t) = (p_1(t), \ldots, p_n(t))$, where

$$p_j(t) = p_j^0 \mathrm{e}^{\left(v_j^0/p_j^0\right)t}, \qquad j = 1, \ldots, n. \tag{25}$$

So, we conclude that (M, G) is also complete. In this case, from (25) and (24), there exist $\alpha_i^k \geq 0$ such that the steepest descent iteration of the sequence generated by the method 1 is given by

$$p_j^{k+1} = p_j^k \mathrm{e}^{\left(v_j^k/p_j^k\right)t_k}, \quad v_j^k = -\sum_{i \in I(p^k, v^k)} \alpha_i^k (p_j^k)^2 \frac{\partial f_i}{\partial p_j}(p^k), \quad \sum_{i \in I(p^k, v^k)} \alpha_i^k = 1, \quad i = 1, \ldots, n.$$

6.2 A steepest descent method for the hypercube

Let M be the hypercube $(0, 1) \times \ldots \times (0, 1)$ endowed with the Riemannian metric

$$M \ni v \mapsto G(p) = P^{-2}(I - P)^{-2} := \operatorname{diag}\left((p_1)^2(1 - p_1)^2, \ldots, (p_n)^2(1 - p_n)^2\right),$$

(metric induzed by the Hessian of the barrier $b(p) = \sum_{i=1}^n (2p_i - 1)\big(\ln p_i - \ln(1 - p_i)\big)$. The Riemannian manifold (M, G) is complete and the geodesic $p = p(t)$, satisfying $p(0) = p^0 = (p_1^0, \ldots, p_n^0)$ and $p'(0) = v^0 0 = (v_1^0, \ldots, v_n^0)$, is given by $p(t) = (p_1(t), \ldots, p_n(t))$,

$$p_j(t) = (1/2)\left[1 + \tanh\left((1/2)\, \frac{v_j}{p_j(1 - p_j)} t + (1/2) \ln\left(\frac{p_j}{1 - p_j}\right)\right)\right], \qquad j = 1, ..., n, \tag{26}$$

where $\tanh(z) := (e^z - e^{-z})/(e^z + e^{-z})$. Moreover (M, G) has constant curvature equal to zero, see Theorem 3.1 and 3.2 of [29]. In this case, from (26) and (24), there exist $\alpha_i^k \geq 0$ such that the steepest descent iteration of the sequence generated by the method 1 is given by

$$p_j^{k+1} = (1/2)\left[1 + \tanh\left((1/2)\,\frac{v_j^k}{p_j^k(1-p_j^k)}t_k + (1/2)\ln\left(\frac{p_j^k}{1-p_j^k}\right)\right)\right], \qquad j = 1, ..., n,$$

with,

$$v_j^k = -\sum_{i\in I(p^k,v^k)} \alpha_i^k (p_j^k)^2(1-p_j^k)^2\frac{\partial f_i}{\partial p_j}(p^k), \qquad \sum_{i\in I(p^k,v^k)} \alpha_i^k = 1, \qquad j = 1, \ldots, n.$$

6.3 steepest descent method for the cone of positive semidefinite matrices

Let $\mathbb{S}^n$ be the set of the symmetric matrices $n \times n$, $\mathbb{S}^n_+$ the cone of the symmetric positive semi-definite matrices and $\mathbb{S}^n_{++}$ the cone of the symmetric positive definite matrices. Following Rothaus [32], let $M = \mathbb{S}^n_{++}$ be endowed with the Riemannian metric induced by the Euclidean Hessian of $\Psi(X) = -\ln\det X$, i.e., $G(X) := \Psi''(X)$. In this case, the unique geodesic segment connecting any $X, Y \in M$ is given by

$$X(t) = X^{1/2}\left(X^{-1/2}YX^{-1/2}\right)^t X^{1/2}, \qquad t \in [0, 1],$$

see [27]. More precisely, M is a Hadamard manifold (with curvature not identically zero), see for example [20], Theorem 1.2, page 325. In particular, the unique geodesic $X = X(t)$ such that $X(0) = X$ and $X'(0) = V$ is given by

$$X(t) = X^{1/2}\mathrm{e}^{tX^{-1/2}VX^{-1/2}}X^{1/2}. \tag{27}$$

Thus, from (27) and (24), there exist $\alpha_j^k \geq 0$ such that the steepest descent iteration of the sequence generated by the method 1 is given by

$$X^{k+1} = (X^k)^{1/2}\mathrm{e}^{t_k(X^k)^{-1/2}V^k(X^k)^{-1/2}}(X^k)^{1/2},$$

with,

$$V^k = -\sum_{i\in I(X^k,V^k)} \alpha_i^k X^k f_i'(X^k)X^k, \qquad \sum_{i\in I(X^k,V^k)} \alpha_i^k = 1.$$

Remark 24. *Under the assumption of convexity on the vector function F, if (M, G) is the Riemannian manifod in the first or in the second example, then Corollary 23 assures the full convergence of the sequence generated by Method 1. This fact doesn't necessarily happens if (M, G) is the Riemanniana manifold in the last example, since in this case (M, G) has curvature not-positive, i.e., $K \leq 0$. However, Theorem 10 assures at least partial convergence.*

7 Final remarks

We have extended the steepest descent method with Armijo's rule for multicriteria optimization to the Riemannian context. Full convergence is obtained under to the assumptions of quasi-convexity of the multicriteria function and non-negative curvature of the Riemannian manifold. A subject in open is to obtain the same result without restrictive assumption on the curvature of the manifold. Following the same line of this paper, as future propose we have the extension, to the context Riemannian, of the proximal method (see Bonnel et al. [5]) and Newton method (see Fliege et al. [13]), both for multiobjective optimization.

References

[1] Absil, P. -A., Baker, C. G., Gallivan, K. A. *Trust-region methods on Riemannian manifolds.* To appear in Foundations of Computational Mathematics. 7 (2007), no.(3), 303-330.

[2] Attouch, H., Bolte, J., Redont, P., Teboulle, M. *Singular Riemannian barrier methods and gradient-projection dynamical systems for constrained optimization.* Optimization. 53 (2004), no. 5-6, 435-454.

[3] Azagra, D., Ferrera, J. López-Mesas, M. *Nonsmooth analysis and Hamilton-Jacobi equations on Riemannian manifolds.* Jornal of Functional Analysis. 220 (2005), 304-361.

[4] Barani, A., Pouryayevali, M.R.,*Invariant monotone vector fields on Riemannian manifolds*, Nonlinear Analysis, Theory, Methods and Applications, 70(5), (2009), 1850-1861.

[5] Bonnel, H., Iusem, A. N, Svaiter, B. F., *Proximal methods in vector optimization.* SIAM J. Optim. 15 (2005), no. 4, 953-970.

[6] Burachik, R., Drummond, L. M. Graña, Iusem, A. N., Svaiter, B. F. *Full convergence of the steepest descent method with inexact line searches.* Optimization 32 (1995), no. 2, 137-146.

[7] da Cruz Neto, J. X., de Lima, L. L., Oliveira, P. R. *Geodesic algorithms in Riemannian geometry.* Balkan J. Geom. Appl. 3 (1998), no. 2, 89-100.

[8] da Cruz Neto, J. X., Ferreira, O. P., Lucâmbio Pérez, L. R., Németh, S. Z. *Convex-and Monotone-Transformable Mathematical Programming Problems and a Proximal-Like Point Method.* Journal of Global Optimization. 35 (2006), 53-69.

[9] do Carmo, M. P. *Riemannian Geometry.* Boston, Birkhauser, (1992).

[10] Ferreira, O. P., *Proximal subgradient and a characterization of Lipschitz function on Riemannian manifolds* J. Math. Anal. Appl. 313, no. 2, (2006), 587-597.

[11] Ferreira, O. P., B. F., Svaiter, *Kantorovich's Theorem on Newton's Method in Riemannian manifolds*. Journal of Complexity. 18 (2002), 304-329.

[12] Ferreira, O. P., Oliveira, P. R. *Subgradient algorithm on Riemannian manifolds*. Journal of Optimization Theory and Applications. 97 (1998), no.1, 93-104.

[13] Fliege, J.; Graa Drummond, L. M.; Svaiter, B. F. *Newton's method for multiobjective optimization*. SIAM J. Optim. 20 (2009), no. 2, 602-626.

[14] Fliege, J., Svaiter, B. F. *Steepest descent methods for multicriteria optimization*. Math. Methods Oper. Res. 51 (2000), no. 3, 479-494.

[15] Gabay, D., *Minimizing a Differentiable Function over a Differentiable Manifold*. Optim. Theory Appl. 37 (1982), pp. 177-219.

[16] Graña Drummond, L. M., Iusem, A. N. *A projected gradient method for vector optimization problems*. Comput. Optim. Appl. 28 (2004), no. 1, 5-29.

[17] Graña Drummond, L. M., Svaiter, B. F. *A steepest descent method for vector optimization*. J. Comput. Appl. Math. 175 (2005), no. 2, 395-414.

[18] Hiriart-Urruty, J.-B, Lemaréchal, C. *Convex analysis and minimization algorithms I and II*, Springer-Verlag, (1993).

[19] Kiwiel, K. C., Murty, K. *Convergence of the steepest descent method for minimizing quasiconvex functions*. J. Optim. Theory Appl. 89 (1996), no. 1, 221-226.

[20] Lang, S. *Fundamentals of Differential Geometry*, Springer - Verlag, (1998).

[21] Ledyaev, Yu. S., Zhu, Qiji J. *Nonsmooth analysis on smooth manifolds*. Trans. Amer. Math. Soc. 359 (2007), no. 8, 3687-3732 (electronic).

[22] Li, C., López, G., Martín-Márquez, V. *Monotone vector fields and the proximal point algorithm on Hadamard manifolds*. J. London Math. Soc., 79(2), 2009, pp.663-683.

[23] Li, S. L., Li, C., Liou, Y. C., Yao, J. C. *Existence of solutions for variational inequalities on Riemannian manifolds*, Nonliear Anal., 71(11), 2009, pp.5695-5706.

[24] Luc, T. D., *Theory of vector optimization*, Lecture Notes in Economics and Mathematical Systems, vol. 319, Springer, Berlin, 1989.

[25] Luenberger, David G. *The gradient projection method along geodesics*. Management Sci. 18 (1972), 620-631.

[26] Munier, J. *Steepest descent method on a Riemannian manifold: the convex case*, Balkan Jour. Geom. Appl. 12, 2 (2007), 98-106.

[27] Nesterov, Y. E. and Todd, M. J. *On the Riemannian Geometry Defined by Self-Concordant Barriers and Interior-Point Methods*, Found. Comput. Math. 2 (2002), no. 4, 333-361.

[28] Papa Quiroz, E. A., Quispe, E. M., Oliveira, P. R. *Steepest descent method with a generalized Armijo search for quasiconvex functions on Riemannian manifolds* J. Math. Anal. Appl. 341, no. 1 (2008), 467-477.

[29] Papa Quiroz, E. A.; Oliveira, P. R. *New Self-Concordant Barrier for the Hypercube* J Optim Theory Appl 135 (2007) 475-490.

[30] Papa Quiroz, E. A.; Oliveira, P. R. *Proximal point methods for quasiconvex and convex functions with Bregman distances on Hadamard manifolds.* J. Convex Anal. 16 (2009), no. 1, 49-69.

[31] Rapcsák, T. *Smooth nonlinear optimization in* R^n. Kluwer Academic Publishers, Dordrecht, (1997).

[32] Rothaus, O. S., *Domains of positivity*, Abh. Math. Sem. Univ. Hamburg. 24 (1960) 189-235.

[33] Sakai, T. *Riemannian geometry.* Translations of mathematical monographs, 149, Amer. Math. Soc., Providence, R.I. (1996).

[34] Smith, S. T. *Optimization techniques on Riemannian Manifolds.* Fields Institute Communications, Amer. Math. Soc., Providence, R.I. 3 (1994), 113-146.

[35] Udriste, C. *Convex functions and optimization methods on Riemannian manifolds.* Mathematics and its Applications. 297, Kluwer Academic Publishers (1994).

[36] J. H. Wang and C. Li, *Convergence of the family of Euler-Halley type methods on Riemannian manifolds under the r-ondition*, Taiwanese J. Math., 13(2), (2009). 585-606.

[37] J. H. Wang and C. Li, *Kantorovich's theorems of Newton's method for mappings and optimization problems on Lie groups*, IMA Numer, Anal,, (2009) to appear.

[38] J. H. Wang, S. C. Huang and C. Li, *Extended Newton's method for mappings on Riemannian manifolds with values in a cone*, Taiwanese J. Math., 13, (2009), 633-656.